PLoS Biology 2017 vol.15 N9

Systematic analysis of transcription start sites in avian development

Lizio M., Deviatiiarov R., Nagai H., Galan L., Arner E., Itoh M., Lassmann T., Kasukawa T., Hasegawa A., Ros M., Hayashizaki Y., Carninci P., Forrest A., Kawaji H., Gusev O., Sheng G.
Kazan Federal University, 420008, Kremlevskaya 18, Kazan, Russia

Abstract

© 2017 Lizio et al. Cap Analysis of Gene Expression (CAGE) in combination with single-molecule sequencing technology allows precision mapping of transcription start sites (TSSs) and genome-wide capture of promoter activities in differentiated and steady state cell populations. Much less is known about whether TSS profiling can characterize diverse and non-steady state cell populations, such as the approximately 400 transitory and heterogeneous cell types that arise during ontogeny of vertebrate animals. To gain such insight, we used the chick model and performed CAGE-based TSS analysis on embryonic samples covering the full 3-week developmental period. In total, 31,863 robust TSS peaks (> 1 tag per million [TPM]) were mapped to the latest chicken genome assembly, of which 34% to 46% were active in any given developmental stage. ZENBU, a web-based, open-source platform, was used for interactive data exploration. TSSs of genes critical for lineage differentiation could be precisely mapped and their activities tracked throughout development, suggesting that non-steady state and heterogeneous cell populations are amenable to CAGE-based transcriptional analysis. Our study also uncovered a large set of extremely stable housekeeping TSSs and many novel stage-specific ones. We furthermore demonstrated that TSS mapping could expedite motif-based promoter analysis for regulatory modules associated with stage-specific and housekeeping genes. Finally, using Brachyury as an example, we provide evidence that precise TSS mapping in combination with Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR)-on technology enables us, for the first time, to efficiently target endogenous avian genes for transcriptional activation. Taken together, our results represent the first report of genome-wide TSS mapping in birds and the first systematic developmental TSS analysis in any amniote species (birds and mammals). By facilitating promoter-based molecular analysis and genetic manipulation, our work also underscores the value of avian models in unravelling the complex regulatory mechanism of cell lineage specification during amniote development.

http://dx.doi.org/10.1371/journal.pbio.2002887

References

[1] Kanamori-Katayama M, Itoh M, Kawaji H, Lassmann T, Katayama S, Kojima M, et al. Unamplified cap analysis of gene expression on a single-molecule sequencer. Genome Res. 2011;21(7):1150–9. doi: 10.1101/gr.115469.110 21596820

[2] Kodzius R, Kojima M, Nishiyori H, Nakamura M, Fukuda S, Tagami M, et al. CAGE: cap analysis of gene expression. Nature methods. 2006;3(3):211–22. doi: 10.1038/nmeth0306-211 16489339

[3] Forrest ARR, Kawaji H, Rehli M, Baillie JK, de Hoon MJL, Haberle V, et al. A promoter-level mammalian expression atlas. Nature. 2014;507(7493):462–+. doi: 10.1038/nature13182 24670764

[4] Andersson R, Gebhard C, Miguel-Escalada I, Hoof I, Bornholdt J, Boyd M, et al. An atlas of active enhancers across human cell types and tissues. Nature. 2014;507(7493):455–461. doi: 10.1038/nature12787 24670763

[5] Arner E, Daub CO, Vitting-Seerup K, Andersson R, Lilje B, Drablos F, et al. Transcribed enhancers lead waves of coordinated transcription in transitioning mammalian cells. Science. 2015;347(6225):1010–4. doi: 10.1126/science.1259418 25678556

[6] Dreos R, Ambrosini G, Groux R, Cavin Perier R, Bucher P, The eukaryotic promoter database in its 30th year: focus on non-vertebrate organisms. Nucleic acids research. 2017;45(D1):D51–D5. doi: 10.1093/nar/gkw1069 27899657

[7] Haberle V, Li N, Hadzhiev Y, Plessy C, Previti C, Nepal C, et al. Two independent transcription initiation codes overlap on vertebrate core promoters. Nature. 2014;507(7492):381–5. doi: 10.1038/nature12974 24531765

[8] Nepal C, Hadzhiev Y, Previti C, Haberle V, Li N, Takahashi H, et al. Dynamic regulation of the transcription initiation landscape at single nucleotide resolution during vertebrate embryogenesis. Genome Res. 2013;23(11):1938–50. doi: 10.1101/gr.153692.112 24002785

[9] Hamburger V, Hamilton HL, A Series of Normal Stages in the Development of the Chick-Embryo, (Reprinted from Journal of Morphology, Vol 88, 1951). Dev Dynam. 1992;195(4):231–&.

[10] Stern CD, The chick: A great model system becomes even greater. Dev Cell. 2005;8(1):9–17. doi: 10.1016/j.devcel.2004.11.018 15621526

[11] Sheng G, Epiblast morphogenesis before gastrulation. Developmental biology. 2015;401(1):17–24. doi: 10.1016/j.ydbio.2014.10.003 25446532

[12] Hillier LW, Miller W, Birney E, Warren W, Hardison RC, Ponting CP, et al. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature. 2004;432(7018):695–716. doi: 10.1038/nature03154 15592404

[13] Schmid M, Smith J, Burt DW, Aken BL, Antin PB, Archibald AL, et al. Third Report on Chicken Genes and Chromosomes 2015. Cytogenet Genome Res. 2015;145(2):78–179. doi: 10.1159/000430927 26282327

[14] Warren WC, Hillier LW, Tomlinson C, Minx P, Kremitzki M, Graves T, et al. A New Chicken Genome Assembly Provides Insight into Avian Genome Structure. G3-Genes Genom Genet. 2017;7(1):109–17.

[15] Shin M, Alev C, Wu YP, Nagai H, Sheng GJ, Activin/TGF-beta signaling regulates Nanog expression in the epiblast during gastrulation. Mech Develop. 2011;128(5–6):268–78.

[16] Alev C, Nakano M, Wu Y, Horiuchi H, Sheng G, Manipulating the avian epiblast and epiblast-derived stem cells. Methods in molecular biology. 2013;1074:151–73. doi: 10.1007/978-1-62703-628-3_12 23975812

[17] Jean C, Aubel P, Soleihavoup C, Bouhallier F, Voisin S, Lavial F, et al. Pluripotent genes in avian stem cells. Development, growth & differentiation. 2013;55(1):41–51.

[18] Lever M, Brand-Saberi B, Theiss C, Neurogenesis, gliogenesis and the developing chicken optic tectum: an immunohistochemical and ultrastructural analysis. Brain structure & function. 2014;219(3):1009–24.

[19] Picard B, Lefaucheur L, Berri C, Duclos MJ, Muscle fibre ontogenesis in farm animal species. Reproduction, nutrition, development. 2002;42(5):415–31. 12537254

[20] Chen W, Lv YT, Zhang HX, Ruan D, Wang S, Lin YC, Developmental specificity in skeletal muscle of late-term avian embryos and its potential manipulation. Poultry science. 2013;92(10):2754–64. doi: 10.3382/ps.2013-03099 24046424

[21] Zhao Y, Li H, Fang S, Kang Y, Wu W, Hao Y, et al. NONCODE 2016: an informative and valuable data source of long non-coding RNAs. Nucleic acids research. 2016;44(D1):D203–8. doi: 10.1093/nar/gkv1252 26586799

[22] Li A, Zhang J, Zhou Z, Wang L, Liu Y, Liu Y, ALDB: a domestic-animal long noncoding RNA database. PLoS ONE. 2015;10(4):e0124003. doi: 10.1371/journal.pone.0124003 25853886

[23] Muret K, Klopp C, Wucher V, Esquerre D, Legeai F, Lecerf F, et al. Long noncoding RNA repertoire in chicken liver and adipose tissue. Genetics, selection, evolution: GSE. 2017;49(1):6. doi: 10.1186/s12711-016-0275-0 28073357

[24] Hebenstreit D, Fang M, Gu M, Charoensawan V, van Oudenaarden A, Teichmann SA, RNA sequencing reveals two major classes of gene expression levels in metazoan cells. Molecular systems biology. 2011;7:497. doi: 10.1038/msb.2011.28 21654674

[25] Severin J, Lizio M, Harshbarger J, Kawaji H, Daub CO, Hayashizaki Y, et al. Interactive visualization and analysis of large-scale sequencing datasets using ZENBU. Nat Biotechnol. 2014;32(3):217–9. doi: 10.1038/nbt.2840 24727769

[26] Lizio M, Harshbarger J, Shimoji H, Severin J, Kasukawa T, Sahin S, et al. Gateways to the FANTOM5 promoter level mammalian expression atlas. Genome biology. 2015;16:22. doi: 10.1186/s13059-014-0560-6 25723102

[27] Zhang G, Li C, Li Q, Li B, Larkin DM, Lee C, et al. Comparative genomics reveals insights into avian genome evolution and adaptation. Science. 2014;346(6215):1311–20. doi: 10.1126/science.1251385 25504712

[28] Seki R, Li C, Fang Q, Hayashi S, Egawa S, Hu J, et al. Functional roles of Aves class-specific cis-regulatory elements on macroevolution of bird-specific features. Nature communications. 2017;8:14229. doi: 10.1038/ncomms14229 28165450

[29] Batut P, Dobin A, Plessy C, Carninci P, Gingeras TR, High-fidelity promoter profiling reveals widespread alternative promoter usage and transposon-driven developmental gene expression. Genome Res. 2013;23(1):169–80. doi: 10.1101/gr.139618.112 22936248

[30] Eisenberg E, Levanon EY, Human housekeeping genes, revisited. Trends in genetics: TIG. 2013;29(10):569–74. doi: 10.1016/j.tig.2013.05.010 23810203

[31] Burt DW, Chicken genome: current status and future opportunities. Genome Res. 2005;15(12):1692–8. doi: 10.1101/gr.4141805 16339367

[32] Abe H, Gemmell NJ, Abundance, arrangement, and function of sequence motifs in the chicken promoters. BMC Genomics. 2014;15:900. doi: 10.1186/1471-2164-15-900 25318583

[33] Hoskins RA, Landolin JM, Brown JB, Sandler JE, Takahashi H, Lassmann T, et al. Genome-wide analysis of promoter architecture in Drosophila melanogaster. Genome Res. 2011;21(2):182–92. doi: 10.1101/gr.112466.110 21177961

[34] Chen J, Tambalo M, Barembaum M, Ranganathan R, Simoes-Costa M, Bronner ME, et al. A systems-level approach reveals new gene regulatory modules in the developing ear. Development. 2017;144(8):1531–43. doi: 10.1242/dev.148494 28264836

[35] Khan MA, Soto-Jimenez LM, Howe T, Streit A, Sosinsky A, Stern CD, Computational tools and resources for prediction and analysis of gene regulatory regions in the chick genome. Genesis. 2013;51(5):311–24. doi: 10.1002/dvg.22375 23355428

[36] Roellig D, Tan-Cabugao J, Esaian S, Bronner ME, Dynamic transcriptional signature and cell fate analysis reveals plasticity of individual neural plate border cells. eLife. 2017;6.

[37] Uchikawa M, Ishida Y, Takemoto T, Kamachi Y, Kondoh H, Functional analysis of chicken Sox2 enhancers highlights an array of diverse regulatory elements that are conserved in mammals. Dev Cell. 2003;4(4):509–19. 12689590

[38] Alev C, Wu Y, Kasukawa T, Jakt LM, Ueda HR, Sheng G, Transcriptomic landscape of the primitive streak. Development. 2010;137(17):2863–74. doi: 10.1242/dev.053462 20667916

[39] Lowe CB, Clarke JA, Baker AJ, Haussler D, Edwards SV, Feather development genes and associated regulatory innovation predate the origin of Dinosauria. Mol Biol Evol. 2015;32(1):23–8. doi: 10.1093/molbev/msu309 25415961

[40] Bailey TL, Johnson J, Grant CE, Noble WS, The MEME Suite. Nucleic acids research. 2015;43(W1):W39–49. doi: 10.1093/nar/gkv416 25953851

[41] Abu-Bonsrah KD, Zhang D, Newgreen DF, CRISPR/Cas9 Targets Chicken Embryonic Somatic Cells In Vitro and In Vivo and generates Phenotypic Abnormalities. Sci Rep. 2016;6:34524. doi: 10.1038/srep34524 27694906

[42] Veron N, Qu Z, Kipen PA, Hirst CE, Marcelle C, CRISPR mediated somatic cell genome engineering in the chicken. Dev Biol. 2015;407(1):68–74. doi: 10.1016/j.ydbio.2015.08.007 26277216

[43] Cheng AW, Wang H, Yang H, Shi L, Katz Y, Theunissen TW, et al. Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. Cell research. 2013;23(10):1163–71. doi: 10.1038/cr.2013.122 23979020

[44] Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol. 2013;31(9):827–32. doi: 10.1038/nbt.2647 23873081

[45] Dubrulle J, McGrew MJ, Pourquie O, FGF signaling controls somite boundary position and regulates segmentation clock control of spatiotemporal Hox gene activation. Cell. 2001;106(2):219–32. 11511349

[46] Kawaji H, Lizio M, Itoh M, Kanamori-Katayama M, Kaiho A, Nishiyori-Sueki H, et al. Comparison of CAGE and RNA-seq transcriptome profiling using clonally amplified and single-molecule next-generation sequencing. Genome Res. 2014;24(4):708–17. doi: 10.1101/gr.156232.113 24676093

[47] Sharma CM, Vogel J, Differential RNA-seq: the approach behind and the biological insight gained. Current opinion in microbiology. 2014;19:97–105. doi: 10.1016/j.mib.2014.06.010 25024085

[48] Innocenti N, Golumbeanu M, Fouquier d'Herouel A, Lacoux C, Bonnin RA, Kennedy SP, et al. Whole-genome mapping of 5' RNA ends in bacteria by tagged sequencing: a comprehensive view in Enterococcus faecalis. Rna. 2015;21(5):1018–30. doi: 10.1261/rna.048470.114 25737579

[49] Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, et al. Real-time DNA sequencing from single polymerase molecules. Science. 2009;323(5910):133–8. doi: 10.1126/science.1162986 19023044

[50] Gilbert LA, Horlbeck MA, Adamson B, Villalta JE, Chen Y, Whitehead EH, et al. Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. Cell. 2014;159(3):647–61. doi: 10.1016/j.cell.2014.09.029 25307932

[51] Liu XS, Wu H, Ji X, Stelzer Y, Wu X, Czauderna S, et al. Editing DNA Methylation in the Mammalian Genome. Cell. 2016;167(1):233–47 e17. doi: 10.1016/j.cell.2016.08.056 27662091

[52] Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. Nature methods. 2015;12(5):453–7. doi: 10.1038/nmeth.3337 25822800

[53] Gong T, Szustakowski JD, DeconRNASeq: a statistical framework for deconvolution of heterogeneous tissue samples based on mRNA-Seq data. Bioinformatics. 2013;29(8):1083–5. doi: 10.1093/bioinformatics/btt090 23428642

[54] Li Y, Xie X, A mixture model for expression deconvolution from RNA-seq in heterogeneous tissues. BMC bioinformatics. 2013;14Suppl 5:S11.

[55] Scialdone A, Tanaka Y, Jawaid W, Moignard V, Wilson NK, Macaulay IC, et al. Resolving early mesoderm diversification through single-cell expression profiling. Nature. 2016;535(7611):289–+. doi: 10.1038/nature18633 27383781

[56] Patel AP, Tirosh I, Trombetta JJ, Shalek AK, Gillespie SM, Wakimoto H, et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. Science. 2014;344(6190):1396–401. doi: 10.1126/science.1254257 24925914

[57] Poulain S, Kato S, Arnaud O, Morlighem JE, Suzuki M, Plessy C, et al. NanoCAGE: A Method for the Analysis of Coding and Noncoding 5'-Capped Transcriptomes. Methods in molecular biology. 2017;1543:57–109. doi: 10.1007/978-1-4939-6716-2_4 28349422

[58] Weng W, Sheng G, Five transcription factors and FGF pathway inhibition efficiently induce erythroid differentiation in the epiblast. Stem Cell Reports. 2014;2(3):262–70. doi: 10.1016/j.stemcr.2014.01.019 24672750

[59] Itoh M, Kojima M, Nagao-Sato S, Saijo E, Lassmann T, Kanamori-Katayama M, et al. Automated workflow for preparation of cDNA for cap analysis of gene expression on a single molecule sequencer. PLoS ONE. 2012;7(1):e30809. doi: 10.1371/journal.pone.0030809 22303458

[60] Takahashi H, Lassmann T, Murata M, Carninci P, 5' end-centered expression profiling using cap-analysis gene expression and next-generation sequencing. Nature protocols. 2012;7(3):542–61. doi: 10.1038/nprot.2012.005 22362160

[61] Robinson MD, McCarthy DJ, Smyth GK, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139–40. doi: 10.1093/bioinformatics/btp616 19910308